

## Structure and Regulation of the Mouse *ing1* Gene

THREE ALTERNATIVE TRANSCRIPTS ENCODE TWO PHD FINGER PROTEINS THAT HAVE OPPOSITE EFFECTS ON p53 FUNCTION\*

(Received for publication, June 22, 1999, and in revised form, August 30, 1999)

Marija Zeremski†, Jason E. Hill§, Serena S. S. Kwek‡, Irina A. Grigorian‡, Katerina V. Gurova‡, Igor V. Garkavtsev¶, Luda Diatchenko§, Eugene V. Koonin||, and Andrei V. Gudkov‡\*\*

From the ‡Department of Molecular Genetics, College of Medicine, University of Illinois, Chicago, Illinois 60607, §CLONTECH Laboratories, Inc., Palo Alto, California 94303, the ¶Department of Functional Genomics, Genome Therapeutics Corp., Waltham, Massachusetts 02154, and ||NCBI, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894

The human *ING1* gene encodes nuclear protein p33<sup>ING1</sup>, previously shown to cooperate with p53 in cell growth control (Garkavtsev, I., Grigorian, I. A., Ossovskaya, V. S., Chernov, M. V., Chumakov, P. M., and Gudkov, A. V. (1998) *Nature* 391, 295-298). p33<sup>ING1</sup> belongs to a small family of proteins from human, mouse, and yeast of approximately the same size that show significant similarity to one another within the C-terminal PHD finger domain and also contain an additional N-terminal region with subtle but reliably detectable sequence conservation. Mouse *ing1* is transcribed from three differently regulated promoters localized within a 4-kilobase pair region of genomic DNA. The resulting transcripts share a long common region encoded by a common exon and differ in their 5'-exon sequences. Two transcripts are translated into the same protein of 185 amino acids, the mouse equivalent of the human p33<sup>ING1</sup>, while the third transcript encodes a longer protein that has 94 additional N-terminal amino acids. Overexpression of the longer protein interferes with the accumulation of p53 protein and activation of p53-responsive promoters after DNA damage. Between the two products of *ing1*, only the longer one forms a complex with p53 detectable by immunoprecipitation. These results indicate that a single gene, *ing1*, encodes both p53-suppressing and p53-activating proteins that are regulated by alternative promoters.

The *ING1* gene was identified as a result of a functional screening of genes, the suppression of which is associated with neoplastic transformation (1). Inhibition of *ING1* expression by antisense RNA promotes anchorage-independent growth in mouse breast epithelial cells, increases the frequency of focus formation in NIH 3T3 cells, and prolongs the life span of diploid human fibroblasts in culture. *ING1* expression is up-regulated in senescent human fibroblasts (2), and ectopic expression of

*ING1* cDNA leads to G<sub>1</sub> arrest or promotes apoptosis in several cell types (3). The accumulated observations indicating *ING1* participation in the negative regulation of cell proliferation, control of cellular aging, and apoptosis have defined *ING1* as a candidate tumor suppressor gene.

We have recently found that the biological effects of *ING1* and p53 are interrelated and require the activity of both genes. Neither of the two genes can, on its own, cause growth inhibition when the other one is suppressed (4). Furthermore, activation of transcription from the p21/WAF1 promoter, a key mechanism of p53-mediated growth control, depends on the expression of *ING1*. A physical association between p33<sup>ING1</sup> and p53 proteins is detected by immunoprecipitation. These results indicated that p33<sup>ING1</sup> is a component of the p53 signaling pathway and that p33<sup>ING1</sup> cooperates with p53 in negative regulation of cell proliferation by modulating p53-dependent transcriptional activation.

Despite the apparent importance of *ING1* in the control of cell proliferation, our knowledge of expression, regulation, and function of this gene remains incomplete. Moreover, GenBank™ contains two *ING1* mRNA sequences differing in their 5'-ends. The origin of these differences is unknown and requires explanation. So far, all of the information about the function of p33<sup>ING1</sup> was obtained from *in vitro* experiments that involved ectopic expression of *ING1* cDNA or its suppression by antisense RNA. The analysis of the structure of the *ING1* gene and its regulation *in vivo* are essential steps toward the understanding of its function and involvement in developmental and physiological processes. This is particularly important due to the cooperation between *ING1* and p53, which suggests that the functioning of the p53 signaling pathway could be dependent on the regulation of *ING1* expression.

In the present work, we used the mouse ortholog of the human *ING1* gene for detailed structural and expression studies, with the goal of subsequently utilizing it as a system for extensive genetic analysis. *ing1* was found to be a highly evolutionarily conserved gene with complex regulation, which involves generation of alternative transcripts initiated from different promoters and translated into proteins that differ in structure and expression patterns. Moreover, these proteins have opposite effects on p53-regulated transcription, indicating that *ing1* encodes two products with the properties of a putative tumor suppressor and a putative oncogene.

### MATERIALS AND METHODS

**Plasmids and Libraries**—Retroviral vector pLXSN, used for the introduction of mouse and human *ING1* cDNA in NMuMG, 10(1), and ConA cells, was provided by A. Dusty Miller (5). Retroviral vector pLXIG was constructed on the basis of pLXSN vector by substituting SV40 promoter and *neo* sequences with the internal ribosome entry site

\* This work was supported by National Institutes of Health Grants CA60730 and CA75179 (to A. V. G.), by a grant from Quark Biotechnology, Inc. (to A. V. G.), and by Department of the Army Grant DAMD17-97-1-7293 (to M. Z.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF177753, AF177754, AF177755, AF177756, AF177757.

\*\* To whom correspondence should be addressed: Dept. of Molecular Genetics (M/C 669), University of Illinois at Chicago, 900 S. Ashland Ave., Chicago, IL 60607. Tel.: 312-413-0349; Fax: 312-996-0683; E-mail: gudkov@uic.edu.

of encephalomyocarditis virus (6) followed by the enhanced green fluorescent protein sequence (7). This vector permits us to translate both the gene of interest (cloned upstream of the internal ribosome entry site) and enhanced green fluorescent protein from a single bicistronic mRNA transcribed from Moloney murine leukemia virus long terminal repeat promoter. A cDNA library from senescent mouse embryonic fibroblasts was constructed using SuperScript system (Life Technologies, Inc.) according to the manufacturer's protocol and cloned into the  $\lambda$  phage  $\lambda$  gt22A vector. Stratagene's 129SVJ mouse genomic library cloned into the  $\lambda$  FIX II vector was used for the isolation of the mouse genomic, *ing1*-containing clones.

**Cell Lines**—Cultures of mouse embryonic fibroblasts were obtained from 10-day-old embryos. 10(1) cell line, a derivative of Balb/c 3T3 cells that spontaneously deleted both p53 alleles (8), was kindly provided by Arnold Levine. Pseudonormal mouse mammary gland epithelial cell line (NMuMG) was obtained from the ATCC collection. Ecotropic retroviral packaging cell line BOSC23 (9) was kindly provided by Warren Pear and David Baltimore (Massachusetts Institute of Technology). The ConA cell line, a derivative of BALB/c 3T3 cells 12-1 (8) with wild type p53, was described earlier (10). It carries the *lacZ* gene encoding  $\beta$ -galactosidase *Escherichia coli* under the control of the p53-dependent promoter and therefore allows monitoring p53 transcriptional activation by a routine X-gal staining. All cells were maintained in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum supplemented with penicillin and streptomycin. For the serum starvation experiments, 10(1) and NMuMG cells were kept with 0.5% fetal bovine serum for 48 h (NMuMG cells) or 36 h (10(1) cells). For the contact inhibition experiments, cells cultures were used 48 h (NMuMG) or 36 h (10(1)) after they became confluent. NMuMG and 10(1) cells were irradiated with 10 grays of  $\gamma$ -radiation and used 24 h after treatment. Mouse embryo fibroblasts were propagated until they reached senescence. Populations of ConA cells, expressing different levels of p37<sup>ING1</sup>, were generated by transduction with pLXIG vector, carrying p37<sup>ING1</sup> cDNA, followed by fluorescence-activated cell sorting of cells with different levels of fluorescence.

**Animals**—Organs and embryos of FVB/N mice were used for the RNA isolation and preparation of whole body sections of embryos using a cryostatic microtome for histoblot hybridization.

**Hybridization Screening of cDNA and Genomic Libraries**—cDNA and genomic library screenings were done according to the standard protocols (11) using <sup>32</sup>P-labeled human and mouse *ing1* cDNA probes, respectively.

**5'- and 3'-RACE**<sup>1</sup>—Alternative 5'-ends of the mouse *ing1* were isolated from mouse spleen and mouse brain cDNAs using the Marathon-ready cDNA kit (CLONTECH), according to the protocol suggested by the manufacturer. AP1 adaptor-specific sense primer, provided by CLONTECH, and the *ing1*-specific antisense primer (5'-CCATCTGACTCAGATCTGGATCTTC-3') were used for PCR. Nested PCR was performed using AP2 adaptor-specific sense primer, provided by CLONTECH, and *ing1*-specific antisense primer (5'-CTGCGGATCAGGCCTCTGGATGC-3'). Precise determination of the 5'- and 3'-ends of the mouse *ing1* transcripts was done using the Marathon-2 cDNA amplification kit based on the new SMART PCR cDNA synthesis technology (CLONTECH). It is based on the recently identified ability of Moloney murine leukemia virus reverse transcriptase to add several nucleotides to the 3' terminus of first-strand cDNA during the reverse transcriptase reaction (12). Briefly, when reverse transcriptase reaches the 5'-end of the mRNA, it switches templates and continues synthesizing the SMART template-switching oligonucleotide. The resulting single-stranded cDNA contains the complete 5'-end of the mRNA as well as the sequence complementary to the template-switching oligonucleotide and is then selectively amplified by PCR. In these experiments, poly(A) RNA preparations isolated from the thymus and testis were used. The following sequences were used for the synthesis of antisense *ing1*-specific primers: 5'-AGGTGTGGTGGGATCGCAACGC-3' (for isoform 1a), 5'-CGCGGGGAGCCAGAGCAGAGAAGGT-3' (isoform 1c), and 5'-GGCGTGGCCTGTCAITGTCGTG-3' (isoform 1b). *ing1*-specific sense primer 5'-GCGTGGTCTTCTGTACCAT-3' was used for the PCR amplification of the 3'-end.

**Sequence Analysis**—Sequencing was done using a Sequenase version 2.0 DNA sequencing kit (U.S. Biochemical Corp.) or done by the University of Chicago Cancer Research Center DNA Sequencing Facility. In all cases, both strands were read using multiple vector-specific or

gene-specific primers. Protein sequence data base searches were performed using the gapped BLASTP program and the PSI-BLAST program that iterates the search using profiles constructed from BLAST hits as queries for subsequent iterations (13). Multiple sequence alignments were constructed using the Gibbs sampling option of the MACAW program (14, 15). GenBank™ accession numbers for the described sequences are AF177753–AF177757.

**Southern, Northern, and Western analyses** were done according to standard protocols. IgG1 mouse monoclonal antibody against human recombinant p33<sup>ING1</sup> (16) was used for the detection of *ing1*-encoded proteins. Anti-p53 monoclonal mouse IgG antibody Ab-1 was obtained from Calbiochem. p21 rabbit polyclonal IgG antibody was obtained from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA).

**Immunoprecipitation Analysis**—Immunoprecipitation of p53 was carried out using 1  $\mu$ g of DO-1 (Santa Cruz Biotechnology) and 1  $\mu$ g of Ab-1 (Calbiochem) antibodies per 0.5 ml of cell lysate. Cell lysate was obtained by transient transfection of the corresponding vectors by LipofectAMINE Plus (Life Technologies, Inc.) into p53-negative Saos-2 cells. Each transfected plate of cells was lysed in 0.5 ml of modified radioimmune precipitation buffer (25 mM Tris, pH 7.4, 125 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate with protease inhibitor). Sepharose-protein A (Amersham Pharmacia Biotech) was used to pull out the antibody-protein complexes, and pellets were washed four times with modified radioimmune precipitation buffer with 0.5% Nonidet P-40. Pellets were boiled in SDS-loading dye and run on 10% SDS-polyacrylamide gel electrophoresis. Western blot was carried out with anti-HA-biotin antibodies, with clone 12CAS (Roche Molecular Biochemicals), or with anti-p53 biotinylated antibodies (Roche Molecular Biochemicals) according to standard protocols.

**In Situ Hybridization on Nitrocellulose ("Histoblotting")**—Whole body sections, 20  $\mu$ m thick, were prepared from frozen embryos embedded in blocks of Tissue-Tek O.C.T. and stored at  $-70^{\circ}\text{C}$ . Sections were placed on nitrocellulose (Schleicher & Schuell) to prepare histoblots as described previously (17). Histoblots were hybridized with <sup>33</sup>P-labeled RNA probes synthesized using Ambion's MAXIScript *in vitro* transcription kit. *ing1*-specific sense and antisense probes were synthesized on the pBLUESCRIPT plasmid with the fragment of *ing1* cDNA corresponding to the PHD finger domain using T3 polymerase for the antisense and T7 polymerase for the sense probe.  $\beta$ -Actin antisense RNA was synthesized using the template provided by Ambion. Histoblots were incubated in prehybridization solution (0.75 mg/ml yeast tRNA, 0.75 mg/ml poly(A), 50% formamide, 0.3 M Tris, pH 8.0, 1 mM EDTA, 5 $\times$  Denhardt's solution, 10% dextran sulfate, 10 mM dithiothreitol) at  $42^{\circ}\text{C}$  for 1–4 h. Hybridization was carried out for 12–24 h at  $42^{\circ}\text{C}$  in the same solution containing  $1\text{--}5 \times 10^7$  cpm of probe/ml. After hybridization, histoblots were washed in 4 $\times$  SSC at room temperature for 20 min, in 2 $\times$  SSC, 0.04  $\mu$ g/ml RNase A at  $37^{\circ}\text{C}$  for 30 min, in 2 $\times$  SSC at  $37^{\circ}\text{C}$  for 30 min, in 1 $\times$  SSC at  $60^{\circ}\text{C}$  for 15–30 min, and finally in 0.1 $\times$  SSC at  $60^{\circ}\text{C}$  for 15–30 min and exposed to x-ray film for 2–7 days.

## RESULTS

**Multiple Transcripts of the Mouse *ing1* Gene Differ in Their 5'-End Sequences**—In order to isolate the mouse ortholog of the *ING1* gene, we screened a cDNA library prepared from senescent mouse embryonic fibroblasts, using human *ING1* as a probe. The choice of the library was determined by the fact that *ING1* is expressed at higher levels in senescent, compared with normal, human fibroblasts (2). As a result, several clones were isolated, and the two longest were sequenced. The clones were identical to each other and highly similar to human *ING1* throughout most of their length, except for the 5'-ends, which were different and not homologous to the human gene (Fig. 1A). This observation could be an indication of alternative splicing of mouse *ing1*; however, it could also be potentially explained by cloning artifacts that occurred during the cDNA library preparation. To determine the structure of 5'-end sequences of *ing1*, we used a 5'-RACE technique for the isolation of cDNA sequences corresponding to the 5'-termini of *ing1* mRNA. cDNA was synthesized from mouse spleen and brain mRNA and ligated to synthetic adaptors. The cDNA was amplified by PCR using a sense primer specific for the adaptor and an antisense primer specific for the common part of *ing1* located close to the divergent region. Two fragments of different size were subsequently cloned, sequenced, and compared with the clones iso-

<sup>1</sup> The abbreviations used are: RACE, rapid amplification of cDNA ends; PCR, polymerase chain reaction; X-gal, 5-bromo-4-chloro-3-indolyl  $\beta$ -D-galactopyranoside.

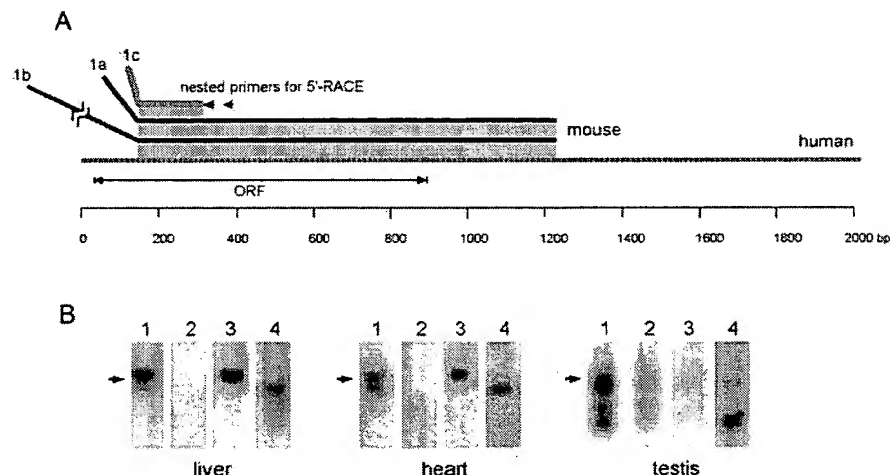


FIG. 1. Cloning of different variants of mouse *ing1* cDNA and identification of their corresponding mRNA transcripts. A, schematic alignment of the mouse *ing1* cDNA clones and 5'-RACE products with human *ING1* cDNA. 1a and 1b represent mouse *ing1* cDNAs isolated from the library of senescent mouse embryonic fibroblasts. A 5'-RACE reaction gave two products, one identical to 1b and another shown as 1c. The position of the gene-specific PCR primers used for the RACE reaction is indicated. Mouse *ing1* sequences are aligned with human *ING1* cDNA. The open reading frame for the human *ING1* clone is indicated. B, analysis of *ing1* transcription by Northern hybridization. *ing1* expression in mouse liver, heart, and testis was analyzed using a multiple-tissue Northern blot (CLONTECH Laboratories), which was hybridized with the probe corresponding to the common part of *ing1* (lane 1), the 5'-end of isoform 1a (lane 2), the 5'-end of isoform 1b (lane 3), or the 5'-end of isoform 1c (lane 4). Probes were obtained by PCR using *ing1*-specific primers. An arrow shows the position of the RNA marker, 2.37 kilobases in size.

lated from the cDNA library. The results of this comparison are schematically presented in Fig. 1A. One of the 5'-RACE products was identical to one of the previously isolated cDNA clones. Another product revealed the third variant of *ing1* cDNA, which again consisted of common and unique regions, with the junction located at exactly the same position as in the other sequences. Fig. 1A shows schematic alignment of the isolated mouse *ing1* cDNA clones and the 5'-RACE products. All variants are identical to each other (and homologous to human *ING1*) up to the same nucleotide and differ only in their 5'-ends.

In order to verify that the observed cDNA structure reflects naturally existing transcripts, we analyzed *ing1* mRNA species by Northern hybridization using probes that represent common or specific regions of the isolated cDNAs. As shown in Fig. 1B, the probe for the common region revealed multiple transcripts in RNA isolated from mouse liver, heart, and testis. Probes specific for individual transcripts, however, showed more simple hybridization patterns, which in combination covered the whole set of transcripts found by hybridization with the common probe. These results indicated that the cloned sequences represent the majority of multiple transcripts of *ing1* synthesized in the thymus and spleen.

**Mapping of Coding Regions of *ing1* in Mouse Genomic DNA**—Comparison of sequences of individual cDNA clones showed differences in their 5' regions, suggesting that *ing1* has multiple alternative 5'-exons (Fig. 1). To verify this hypothesis, we determined the structure of the mouse *ing1* gene. We isolated phage clones carrying sequences homologous to the mouse *ing1* cDNA by hybridization screening of a mouse genomic library. These clones were mapped by restriction digestion analysis, in combination with Southern blot hybridization, with the probes corresponding to the different *ing1* parts. The interpretation of the results obtained is shown in Fig. 2A.

Comparison of genomic and cDNA sequences of *ing1* using PCR with different *ing1*-specific primers (data not shown) as well as with the sequencing data showed that most of the transcribed sequences of the *ing1* gene come from a single exon. Alternative 5'-ends are encoded by different exons positioned upstream from the common exon. Comparison of the sequences of cDNA and genomic clones, as well as 5'-RACE products,

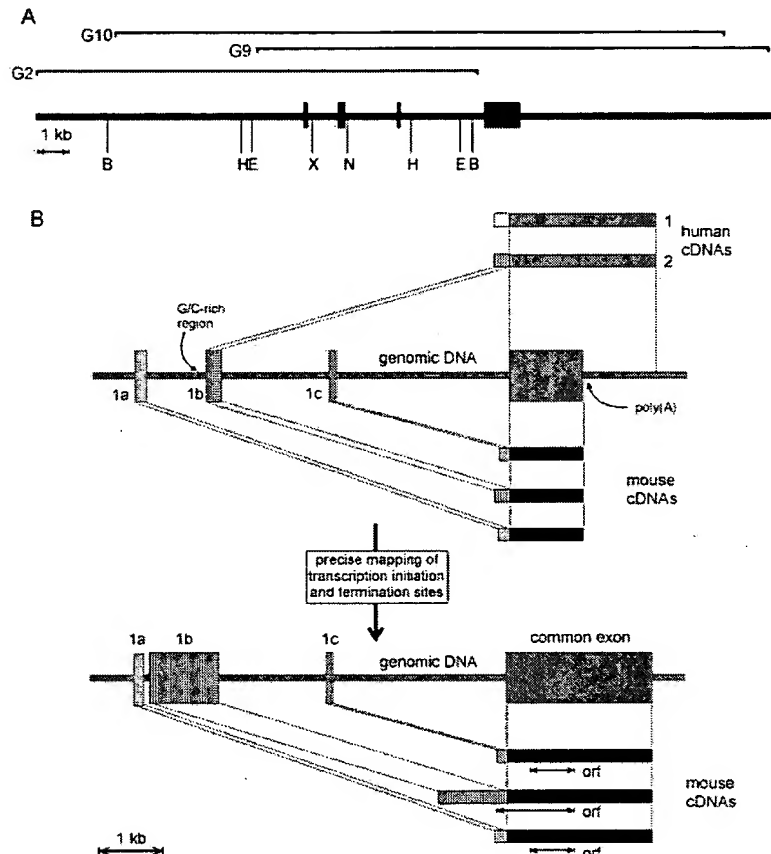
revealed three isoforms of mouse *ing1* that differ from each other only in their 5'-ends, which also indicates that each isoform is most probably expressed from its own promoter. Southern blot hybridization analysis of *ing1*-related sequences in the mouse genome indicated that *ing1* is a single gene with no obvious close family members (data not shown).

The length of the isolated cDNA clones appeared to be significantly shorter than that of the mRNA species detected by Northern hybridization (Fig. 1B), suggesting that part of the transcribed sequences were missing from the isolated cDNAs. To determine the exact start sites of *ing1* transcription, we used a new procedure called "SMART-based 5'-RACE" as described under "Materials and Methods." Using this method, we were able to extend the cDNA sequences for the 5'-end of isoform 1b. This exon contains an extremely G/C-rich region that blocked DNA elongation during the original 5'-RACE reaction (Figs. 2B and 3) and complicated sequencing of the final, long 5'-RACE product. Therefore, the transcription start site of isoform 1b was not precisely identified, since it was estimated based on the size of the 5'-RACE product and the analysis of the genomic sequence.

Comparing the 3'-ends of human and mouse *ING1* cDNAs showed that although the mouse transcript was flanked by a poly(A) stretch, its 3'-untranslated region was significantly shorter than the human one. Analysis of the genomic clone with the mouse *ing1* sequence revealed the presence of a long poly(A) stretch that could potentially be used as a primer-binding sequence for reverse transcription initiated from oligo(dT) primers (Fig. 3C). Moreover, the alignment with the 3'-end of the human *ING1* cDNA sequence resumes downstream from this genomic poly(A) stretch. The exact end of *ing1* transcription in mice was determined using the same method applied for the generation of 5'-sequences. Mouse *ing1* cDNA was, in fact, found to be longer than was originally thought (Figs. 2B and 3). Sequence comparison showed that mouse and human *ING1* transcripts terminate at the same point and share significant levels of similarity up to the very 3'-ends.

**Sequence Analysis of *ing1***—Nucleotide and predicted amino acid sequences of the three mouse *ing1* cDNA isoforms and the promoter regions for all three alternative transcripts are shown in Fig. 3. It was found that *ing1* isoforms 1a and 1c have several

**FIG. 2. Mapping of mRNA-coding regions within the mouse *ing1* locus.** **A**, map of the mouse genomic region containing *ing1* cDNA-related sequences. Three *ing1*-positive genomic clones were isolated from the mouse genomic library using *ing1* cDNA as a probe (G2, G9, and G10), and analyzed by restriction digestion and Southern blot hybridization. This analysis, in combination with the PCR data and with partial sequencing of the genomic clones, allowed us to determine the positions of the *ing1* exons. **B**, determination of transcription initiation and termination sites in the *ing1* gene. The upper panel shows the genomic map with the positions of the *ing1* exons and the structure of *ing1* cDNA clones as determined after cDNA library screening and 5'-RACE analysis. The structure of two known human *ING1* cDNA isoforms in relation to mouse *ing1* sequences is shown above the genomic DNA. The lower panel demonstrates the final structure of the *ing1* gene determined as a result of precise mapping of transcription initiation and termination sites. In the upper panel, the position of the G/C-rich region in the 5'-end of clone 1b and the poly(A) region in the 3' part of the *ing1* transcripts are indicated, which interfered with the polymerase chain reaction during the original 5'-RACE.



tightly clustered transcription start sites. Sequences upstream of the initiation sites (putative promoters) lack TATA boxes, but they do contain a sequence corresponding to the loose initiator consensus PyPyAN(T/A)PyPy that includes the transcription start sites (Fig. 3). In both cases, areas upstream from the initiator are extremely GC-rich with multiple Sp1-binding sites. All of these features are typical of TATA-less promoters (18, 19). Binding sites of inducible factors that are usually present further upstream in the promoter area (e.g. CAAT box, Oct-1 and Oct-2 binding site, NF- $\kappa$ B, or ATF binding site) were not found in these promoters. Transcription of *ing1* isoform 1b is estimated to start approximately 30 nucleotides downstream from the TATA-like box. A putative CAAT box is present 60 nucleotides upstream from the TATA-like sequence. Since the transcription initiation sites for isoforms 1a and 1b are only about 200 base pairs apart, there is a possibility that their promoters share some of the regulatory sequences including, for example, a number of Sp1 binding sites located upstream of this area.

All three alternative transcripts of mouse *ing1* contain the same long open reading frame, although the sizes of the predicted proteins are different. While isoform 1b encodes a protein of 279 amino acids, the other two isoforms are predicted to encode a shorter protein product of 185 amino acids, which lacks 94 N-terminal amino acids. Translation of these two products is expected to start from the initiation codon, which is located at the beginning of the common exon. Alignment of the predicted amino acid sequences with that of human p33<sup>ING1</sup> (GenBank<sup>TM</sup> accession number AF044076) revealed high similarity between the mouse and human proteins (89% sequence identity) (Fig. 4).

The nonredundant protein sequence data base at NCBI was searched using the *ing1*-encoded protein sequence as a query,

and a highly significant sequence similarity was detected with a human paralog of p33<sup>ING1</sup>. These included three uncharacterized proteins from the budding yeast *Saccharomyces cerevisiae* and their homologue from fission yeast *Schizosaccharomyces pombe* (probability of the similarity being observed by chance  $<10^{-12}$ ). All of these proteins are approximately the same size and contain a C-terminal PHD finger domain (20, 21). The sequence conservation in the PHD domain in these six protein sequences is striking. There are 13 invariant residues in addition to the eight metal-chelating cysteines and histidines that are conserved in all PHD fingers (Fig. 4B). Multiple alignment analysis resulted in the delineation of an additional N-terminal region that is conserved in these proteins (Fig. 4B). This region consists of approximately 100 amino acid residues, includes two distinct conserved motifs, and shows a fairly subtle similarity that was not statistically significant in the context of the screening of the complete data base, except for the conservation between the two human paralogs. However, in the reduced search space defined by the presence of the PHD finger, it was shown that the probability of finding this level of similarity by chance was less than  $10^{-13}$  for the distal motif and  $10^{-5}$  for the proximal motif.

**Expression of *ing1* Transcripts in Adult and Embryonic Mouse Tissues**—We analyzed the expression of *ing1* mRNA in the organs of adult mice and mouse embryos at different stages of development by Northern blot hybridization. The probe representing the common exon of the *ing1* gene revealed multiple mRNA size classes that represent alternative transcripts of *ing1* (Fig. 5A). Overall *ing1* mRNA expression is most abundant in thymus and testis. Much lower levels were detected in the rest of the tissues tested, and they also differed in the content and relative intensity of the hybridizing bands. The same pattern of *ing1* mRNA expression was observed in p53-

FIG. 3. Sequences of the putative promoter areas of the *ing1* gene, *ing1* cDNAs, and their predicted protein products. A, sequence of the genomic region with the putative promoter areas of cDNA isoforms 1a and 1b; the start sites of transcription of these two variants are marked by arrows (the start site of isoform 1b was estimated based on the size of the 5'-RACE product and the analysis of genomic sequence). The initiator sequence of isoform 1a, including multiple initiation start sites, is underlined; TATA-like sequence for isoform 1b is shown in **boldface type**. The underlined boldface italic sequence is the CAAT box positioned about 100 base pairs upstream from the 1b transcript initiation site. Sp1 binding sites are displayed in **boldface italic type**. B, sequence of the putative promoter area of isoform 1c, with the underlined initiator sequence





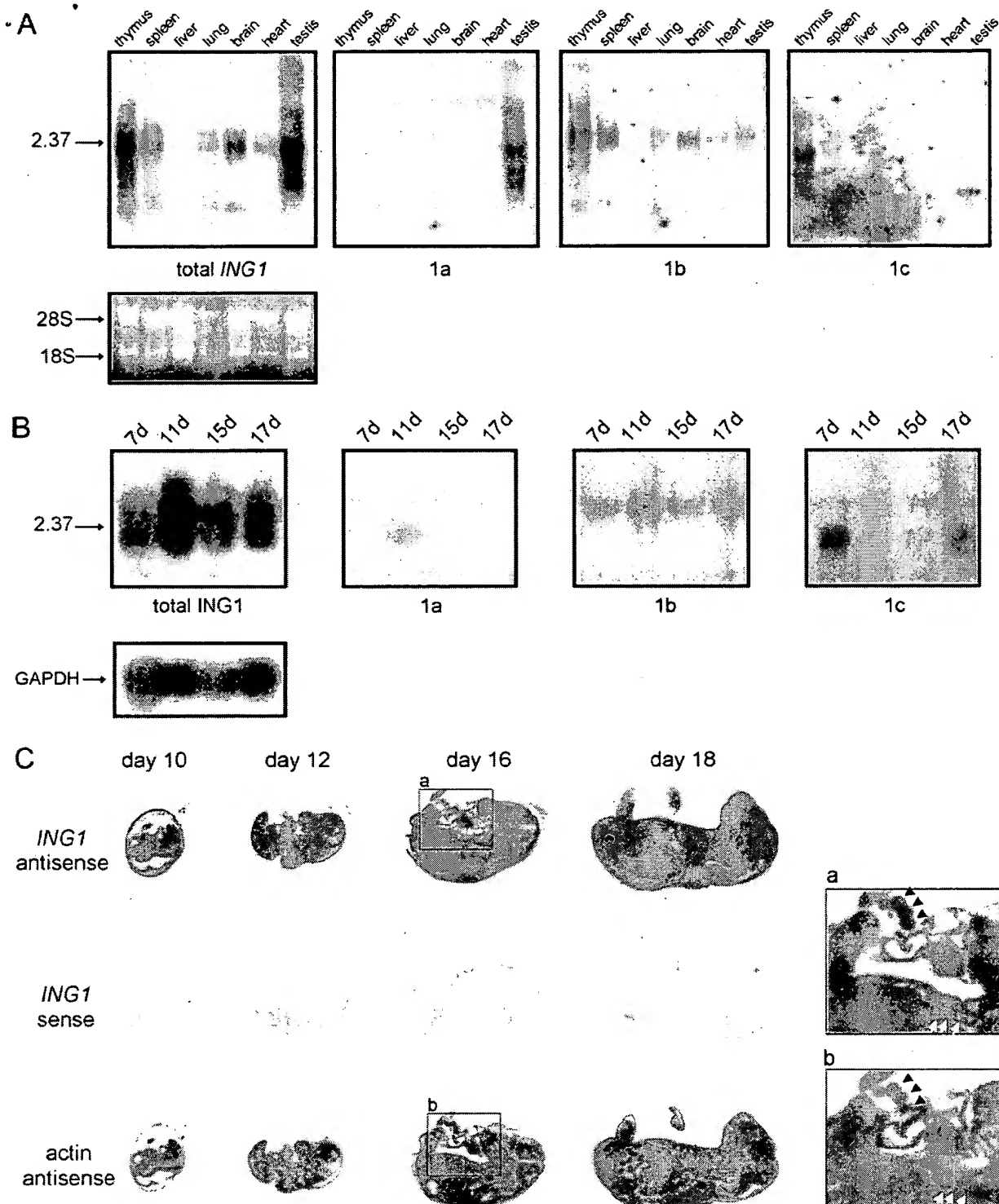
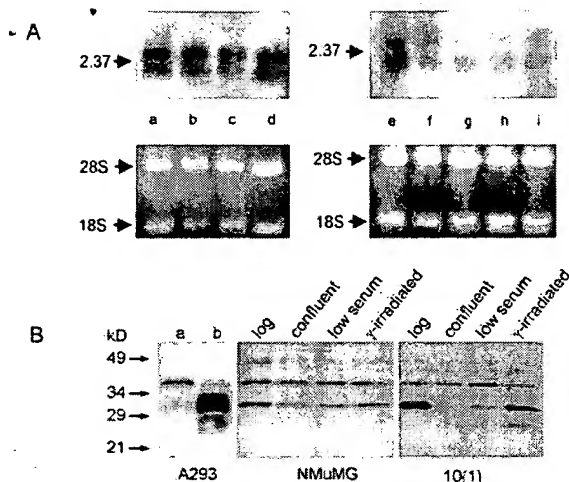


FIG. 5. *ing1* expression in adult mouse tissues and embryos. **A**, results of Northern hybridization of total mouse RNAs isolated from the indicated organs with the indicated probes; a photograph of an ethidium bromide-stained gel is shown as a loading control. **B**, Northern hybridization of RNA on a CLONTECH mouse embryonic multiple-tissue Northern blot (mRNA samples from mouse embryos at 7, 11, 15, and 17 days of development) with the same probes. The arrows indicate positions of the 2.37-kilobase RNA marker as well as the positions of 28 and 18 S rRNAs. **C**, *ing1* expression in mouse embryos analyzed by *in situ* histoblot hybridization. Histoblots were prepared from whole body sections of 10-, 12-, 16-, and 18-day embryos and hybridized with the sense and antisense *ing1*-specific  $^{32}$ P-labeled RNA probes corresponding to the region encoding PHD domain. Antisense probe for ubiquitously expressed actin gene was used as a control.

analyzed, mRNA for isoform 1a was expressed only in testis and also in the 11-day embryo. While there is no detectable signal with the 1a-specific probe in the 7-day embryo, traces of hybridization could be detected in mRNA from day 15 and day

17 embryos, which indicates an extremely low expression level at these points of mouse embryogenesis.

**Indications of Proliferation-dependent Regulation of *ing1* Expression**—To check whether *ing1* expression was proliferation-



**FIG. 6. *ing1* mRNA and protein expression varies depending on cell growth conditions.** A, RNA was isolated from dividing 10(1) (lane a) and NMuMG (lane e) cell lines, as well as from the dividing young mouse embryonic fibroblasts (lane h). RNA was also isolated from contact-inhibited 10(1) (lane b) and NMuMG cells (lane f); serum-starved 10(1) cells (lane c) (64 h at 0.5% FBS);  $\gamma$ -irradiated 10(1) cells (lane d) and NMuMG cells (lane g); and senescent MEFs (lane i). Filters were probed with total *ing1* probe. The bottom panels show RNA gels from which Northern blots were made. The arrows indicate positions of the 2.37-kilobase RNA marker (upper panels) or 28 and 18 S rRNAs. B, detection of *ing1*-encoded proteins by Western immunoblotting, using antibodies against p33<sup>ING1</sup>. Cell lysates were prepared using radioimmune precipitation buffer from the cells treated in the same way as explained above. As a control, cell lysates were prepared from 293 cells transfected with the plasmids expressing mouse 1b *ing1* variant (lane a) or the human homologue of p31<sup>ING1</sup> that produces the protein product of the same size as mouse 1c and 1a variants.

dependent, we analyzed RNA from two mouse cell lines, NMuMG and 10(1), at different growth conditions by Northern blot hybridization using total *ing1* cDNA as a probe. *ing1* expression was also compared in senescent versus young, dividing mouse embryonic fibroblasts. Results are shown in Fig. 6A. In both cell lines, *ing1* was expressed at higher levels in dividing compared with quiescent cells (quiescence is induced by serum starvation, contact inhibition, or  $\gamma$ -irradiation). However, in 10(1) cells the difference in expression was specific only for the upper *ing1* specific band, which corresponds to isoform 1b. In mouse embryonic fibroblasts, *ing1* was expressed at very low levels without any differences between dividing and senescent cells.

*ing1* expression in cell lines was also analyzed by Western blotting using an IgG1 mouse monoclonal antibody produced against human recombinant p33<sup>ING1</sup> (16) (Fig. 6B). This antibody has previously been shown to specifically detect the denatured form of mouse p33<sup>ING1</sup> protein in Western immunoblot protocols. Cell lysates were produced from dividing, contact-inhibited, serum-starved, or  $\gamma$ -irradiated 10(1) and NMuMG cells. In both cell lines, two protein products were detected, 31 and 37 kDa in size, which correspond to the truncated *ING1* and 1b *ing1* protein products, respectively. The 37-kDa protein was present at the same levels in dividing and quiescent cells, while the 31-kDa protein was present at higher levels in dividing compared with nondividing cells. The variation of expression of *ing1* differs significantly at mRNA and protein levels, indicating that this gene could be regulated at the level of transcription or protein stability.

**Overexpression of p37<sup>ING1</sup> Inhibits p53 Function**—All previous functional analysis of *ING1* was done using human cDNA expressing the shorter protein encoded by this gene (1, 2, 4). As a result, p33<sup>ING1</sup> was defined as a negative growth regulator

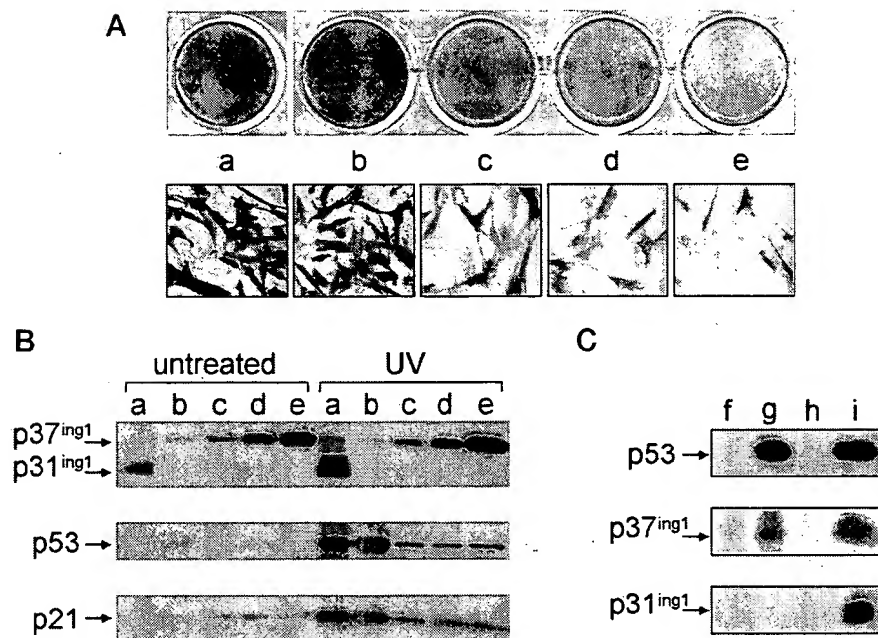
that cooperates with p53 in transcriptional activation of p53-responsive genes. Its mouse homologue, p31<sup>ING1</sup>, has a similar biological effect (data not shown) that is consistent with the high degree of identity between the mouse and the human proteins (see Fig. 4). To analyze the influence of the longer product of *ing1* on p53 function, we overexpressed cDNA for p37<sup>ING1</sup> in mouse cells maintaining wild type p53. To effectively monitor p53-dependent transcriptional activation, we used previously characterized ConA cells (see "Materials and Methods"), carrying the p53-responsive bacterial *lacZ* gene. We transduced ConA cells with retrovirus expressing p37<sup>ING1</sup> and enhanced green fluorescent protein from a bicistronic mRNA and generated cell populations with different levels of overexpression of p37<sup>ING1</sup> by sorting cells differing in the levels of enhanced green fluorescent protein fluorescence. These cells were compared with ConA cells transduced with either the empty pLXIG vector or with retroviral vector expressing the short protein product of *ing1*. The expression of introduced constructs was confirmed by Western immunoblotting using an anti-p33<sup>ING1</sup> antibody that recognizes both human and mouse proteins (16) (Fig. 7B). p53-dependent expression of *lacZ* was determined by X-gal staining of ConA cells treated with UV light (Fig. 7A). The expression of p53 protein and p53-responsive p21<sup>waf1</sup> protein in untreated and in UV-irradiated cells was estimated by Western immunoblotting (Fig. 7B). A decreased intensity of the p21 band by Western blot and a weaker X-gal staining of ConA cells treated with UV light indicate that p37<sup>ING1</sup> clearly inhibited UV-induced accumulation of both p53-inducible proteins. Moreover, overexpression of p37<sup>ING1</sup> strongly attenuated accumulation of p53 in response to UV light (Fig. 7B). These observations demonstrate that p37<sup>ING1</sup> has a suppressive effect on p53-dependent transcriptional activation supposedly by attenuating p53 accumulation in response to DNA damage. This means that the two proteins encoded by the *ing1* gene have opposite effects on p53 function.

It was previously shown that p33<sup>ING1</sup> can be detected in cells in a complex with p53 by co-immunoprecipitation (4). Since at that time it was not known that *ING1* encodes two different proteins, it was unclear which of the two products of *ING1* was detected in these experiments. We therefore compared the ability of short and long *ING1*-encoded proteins to form a complex with p53 by using immunoprecipitation. Since the protocol of immunoprecipitation was optimized for human proteins, we analyzed human variants of *ING1* proteins expressed in human cells by transient transfection. p53-deficient Saos-2 cells were used in these experiments that have both alleles of the p53 gene deleted and presumably retain an intact p53 pathway. They were co-transfected in different combinations with the expression plasmids encoding wild type p53 and the short or long products of *ING1*. Cell lysates were treated with anti-p53 antibodies (see "Materials and Methods"), and the presence of *ING1*-encoded proteins in precipitates was detected by Western immunoblotting (Fig. 7C). The results obtained showed that although p37<sup>ING1</sup> is co-precipitated with p53, the shorter *ING1* product is not detectable in the p53 precipitate. This observation suggests that differences in the effect of the two *ING1* products on p53 function could be a reflection of differences in their interaction with p53 in the cell.

#### DISCUSSION

**Multiple Products of the *ing1* Gene**—*ING1* was originally described as a gene whose suppression promotes neoplastic transformation (1). Consistent with that, ectopic expression of the first isolated human cDNA was growth-suppressive for different cell lines. *ING1* was found to encode a nuclear protein termed p33<sup>ING1</sup> that was shown to functionally and physically interact with p53 in cell growth regulation (4, 22). Now, after





**FIG. 7. p37<sup>ing1</sup> affects p53 function.** A, dependence of  $\beta$ -galactosidase activation after UV irradiation of ConA cells on the expression levels of p37<sup>ing1</sup>. ConA cells expressing the human homologue of p31<sup>ing1</sup> (lane a); p37<sup>ing1</sup> at low (lane c), medium (lane d), or high (lane e) levels; and ConA cells infected with pLXIG vector alone (lane b) were equally plated in a 12-well plate, UV-irradiated (25 J/m<sup>2</sup>), and X-gal-stained 16 h after the treatment. A fragment of the 12-well plate and microscopic views of individual wells are shown. B, expression of p53 protein, p53-responsive p21<sup>ras</sup> protein, and *ing1* protein products in untreated and in UV-irradiated ConA cells was estimated by Western immunoblotting. Analyzed cell lysates were isolated using radioimmune precipitation buffer from the cells treated in the same way as explained for A. C, coimmunoprecipitation of p53 and *ING1* protein products in Saos-2 cells. Detection of p53 (upper panel), p37<sup>ING1</sup> (middle panel), and p31<sup>ING1</sup> (lower panel) proteins by Western immunoblotting is shown. Lanes f, g, and i (top and middle panels), lysates from Saos-2 cells co-transfected with p53 and p37<sup>ING1</sup>-expressing plasmids. Lanes f, g, and i (bottom panel), lysates from Saos-2 cells co-transfected with p53 and p31<sup>ING1</sup>-expressing plasmids. Lane h, lysate from control nontransfected cells. Lanes f and g contain the products of immunoprecipitation with anti-p53 antibodies (lane g) and control nonspecific antibodies (lane f).

we characterized the structure and expression of the mouse *ing1* gene, it became clear that *ing1* regulation is more complicated than was originally thought. *ing1* is transcribed from at least three differently regulated promoters, and the resulting transcripts encode at least two different proteins.

A similar type of regulation (alternative initiation leading to variability of 5'-exons) found in other tumor suppressor genes, namely *BRCA1*, *APC*, and *INK4* (23–26), is associated with the generation of proteins with different functions (in the case of *INK4*, different reading frames are translated in the alternative transcripts). Consistently, the two products of *ing1* have opposite effects on p53-dependent transcription regulation; one acts as a p53 cooperator, while the other acts as a p53 suppressor.

All of the transcripts share a common region encoded by a common exon but differ in their 5'-exons. Two of these alternative exons do not contain protein-coding sequences (isoforms 1a and 1c), while the third one does (isoform 1b). Consistently, one of the *ing1* transcripts encodes a 37-kDa protein (p37<sup>ING1</sup>), while two others are translated into a shorter protein of 24 kDa that surprisingly runs as if it was 31 kDa (p31<sup>ING1</sup>). Structures of mouse and human *ing1* genes are likely to be similar, considering the high degree of evolutionary conservation of *ing1* sequences and the fact that the two versions of human *ING1* that are currently available in GenBank<sup>TM</sup> also share a large common part but have different N termini. It is noteworthy that one of the human variants of *ING1* (GenBank<sup>TM</sup> accession number AF001954) does not have a homologue with similar 5'-sequences among the identified mouse *ing1* isoforms, which raises the possibility that there could be more variants of mouse *ing1* transcripts that so far have not been isolated.

p33<sup>ING1</sup> and its homologs in such a distantly related species

as yeast contain a remarkably conserved PHD finger domain and an additional, weakly conserved domain of unknown function. PHD fingers have been shown to bind DNA (21, 27), but to our knowledge there is no evidence that they mediate protein-protein interactions. Therefore, it seems likely that the C-terminal PHD domain in the p33<sup>ING</sup> family of proteins is involved in specific DNA binding that may be important for transcription regulation. Given the outstanding conservation of this domain across the large phylogenetic distance that separates humans and yeast, the specificity of DNA binding with respect to particular binding sites is expected to be conserved either in terms of DNA sequence, distinct features of chromatin structure, or both. It may be further surmised that the conserved N-terminal domain is involved in specific protein-protein interactions that couple transcription regulation by p33<sup>ING1</sup> and its homologs to other elements of cell cycle control. The conservation of domain organization in p33<sup>ING1</sup> and its yeast homologs is particularly notable given that yeast does not encode any homologs of p53. Thus, it appears that p33<sup>ING1</sup> is an ancient cell cycle regulator whose interaction with p53 is a later evolutionary addition.

**Regulation of *ing1* Expression**—The results of expression analysis indicated that mouse *ing1* is a subject of regulation both at the protein and mRNA levels and that different isoforms have different expression patterns. Thus, p37<sup>ING1</sup> is ubiquitously expressed in all tissues analyzed with elevated mRNA expression levels in the thymus, while levels of the p31<sup>ING1</sup> protein vary dramatically among organs and cell lines.

Sizes of the *ing1* mRNA transcripts detected by Northern hybridization correlate well with the length of cDNA sequences isolated from all of the tissues except testis. Probes specific for each of the identified splice variants reveal multiple RNA

bands, indicating a variability of the *ing1* transcripts in this organ. It is likely that in testis *ing1* mRNAs are either terminated or processed differently from the rest of the tissues.

There is a clear link between the proliferation rate of the cell and *ing1* expression both at the protein and mRNA levels. This observation may reflect cell cycle dependence of *ing1* transcription detected earlier for the human *ING1* gene; the biological significance of this regulation is not obvious so far. We failed to observe up-regulation of mouse *ing1* in senescent cells as it has been reported for human *ING1* (8), which suggests that regulation patterns of mouse and human *ing1* genes may not always be the same.

***ING1* and p53: Cooperators or Antagonists?**—Human *ING1* has been known as a cooperator of p53 involved in negative growth regulation (1, 4), apoptosis (3), and senescence (2). This set of properties, together with the observations that show a decreased expression of *ING1* in several breast carcinoma cell lines and *ING1* gene rearrangement in one neuroblastoma cell line, allow *ING1* to be defined as a candidate tumor suppressor (1–4, 16, 22). Here we demonstrate that *ING1* encodes, in addition to the protein with the above described properties, another product that shows an opposite effect on p53 function, inhibiting p53-dependent transcriptional activation. Thus, a single *ING1* gene appeared to simultaneously encode a candidate tumor suppressor (p53 cooperator) and a putative oncogene (p53 inhibitor). This means that a simplified view of the potential role of *ING1* in cancer should be revised. In the case of *ING1*, we may have a new genetic mechanism of promoting cancer that involves an imbalance between the two products of one gene. Further characterization of the biological activity of the long product of *ING1*, p37<sup>ING1</sup>, indicated that it can affect the whole spectrum of cell properties known to be controlled by p53, including sensitivity to apoptosis, replicative senescence, cooperation with dominant oncogenes, and drug response.<sup>2</sup> Thus, overexpression of the longer product of *ING1*, which functionally acts as a p53 suppressor, could be the mechanism of attenuation of p53 activity in tumors that do not require mutations in p53 itself.

A large scale analysis of clinical samples is needed to conclude whether this hypothesis is correct or not. However, any attempts to investigate the potential cancer relevance of *ING1* should involve separate analysis of the alternative isoforms of this gene. It has already become obvious that a direct measurement of the overall mRNA expression in tumors cannot be used as an approach to judge the involvement of *ING1* in carcinogenesis. The ratio between expression levels of the two *ING1* products seems to provide much more meaningful information. It is therefore essential to establish such assays that would distinguish between the two protein products or among the alternative *ING1* transcripts.

Although the mechanism of p53 modulation by the two products of the *ING1* gene remains unknown, we hope that the two new observations described in this work will shed light on this problem. We found that (i) p53 is found in a complex with the long but not with the short product of *ING1* and (ii) overexpression of the longer product inhibits accumulation of p53 protein after DNA damage. Both properties of p37<sup>ING1</sup> clearly

resemble those of the natural p53 antagonist, Mdm2, which was shown to affect nuclear localization of p53 (28, 29) and to promote its proteasomal degradation (30, 31). All of these possibilities should be tested for p37<sup>ING1</sup> as well as a potential interdependence of the biological activity of these two proteins.

Is the function of *ING1* limited to its interaction with the p53 signaling pathway? It seems doubtful, considering the high evolutionary conservation of *ING1*. In fact, *ING1* homologues found in yeast (Fig. 4), which does not have detectable p53-related genes, suggest some important p53-independent role for *ING1*. The phenotypes associated with the knockout of *ING1* in yeast and mice may provide a lead toward its currently unknown cellular function. It is noteworthy that a gene knock-out approach will be complicated by the necessity to separately inactivate expression of each of the two products of *ING1*.

**Acknowledgments**—We thank Ilya Mazo for the mouse cDNA library and Elena Komarova for providing histoblots of mouse embryos for *in situ* hybridization.

#### REFERENCES

- Garkavtsev, I., Kazarov, A., Gudkov, A., and Riabowol, K. (1996) *Nat. Genet.* **14**, 415–420
- Garkavtsev, I., and Riabowol, K. (1997) *Mol. Cell. Biol.* **17**, 2014–2019
- Helbing, C. C., Veillette, C., Riabowol, K., Johnston, R. N., and Garkavtsev, I. (1997) *Cancer Res.* **57**, 1255–1258
- Garkavtsev, I., Grigorian, I. A., Ossovskaya, V. S., Chernov, M. V., Chumakov, P. M., and Gudkov, A. V. (1998) *Nature* **391**, 295–298
- Miller, A. D., and Rosman, G. J. (1989) *BioTechniques* **7**, 980–986
- Jang, S. K., Krausslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., and Wimmer, E. (1988) *J. Virol.* **62**, 2636–2643
- Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) *Gene (Amst.)* **173**, 33–38
- Harvey, D. M., and Levine, A. J. (1991) *Gen. Dev.* **5**, 2375–2385
- Pear, W. S., Nolan, G. P., Scott, M. L., and Baltimore, D. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8392–8396
- Komarova, E. A., Chernov, M. V., Franks, R., Wang, K., Armin, G., Zelnick, C. R., Chin, D. M., Bacus, S. S., Stark, G. R., and Gudkov, A. V. (1997) *EMBO J.* **16**, 1391–1400
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., pp. 8.1–9.62, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Chenichik, A., Zhu, Y. Y., Diatchenko, L., Li, R., Hill, J., and Siebert, P. D. (1998) *Gene Cloning and Analysis by RT-PCR* (Siebert, P. D., and Larrick, J. W., eds) pp. 305–319, BioTechniques Books, Eaton Publishing Co., Natick, MA
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
- Neuwald, A. F., Liu, J. S., and Lawrence, C. E. (1995) *Protein Sci.* **8**, 1618–1632
- Schuler, G. D., Altschul, S. F., and Lipman, D. J. (1991) *Proteins* **9**, 180–190
- Garkavtsev, I., Boland, D., Mai, J., Wilson, H., Veillette, C., and Riabowol, K. (1997) *Hybridoma* **16**, 537–540
- Gudkov, A. V., Kashkin, K. N., Zaitsevskaia, T. E., and Troyanovsky, S. M. (1989) *Int. J. Cancer* **44**, 1052–1056
- Smale, S. T. (1994) *Transcription: Mechanisms and Regulation* (Conaway, R. C., and Conaway, J. W., eds) pp. 63–80, Raven Press, New York
- Smale, S. T. (1997) *Biochim. Biophys. Acta* **1351**, 73–88
- Aasland, R., Gibson, T. J., and Stewart, A. F. (1995) *Trends Biochem. Sci.* **20**, 56–59
- Schindler, U., Beckmann, H., and Cashmore, A. R. (1993) *Plant J.* **4**, 137–150
- Garkavtsev, I., Demetrick, D., and Riabowol, K. (1997) *Cytogenet. Cell Genet.* **76**, 176–178
- Cui, J. Q., Wang, H., Reddy, E. S., and Rao, V. N. (1998) *Oncol. Rep.* **5**, 585–589
- Quelle, D. E., Zindy, F., Ashmun, R. A., and Sherr, C. J. (1995) *Cell* **83**, 993–1000
- Thliveris, A., Samowitz, W., Matsunami, N., Groden, J., and White, R. (1994) *Cancer Res.* **54**, 2991–2995
- Xu, C.-F., Brown, M. A., Chambers, J. A., Griffiths, B., Nicolai, H., and Solomon, E. (1995) *Hum. Mol. Genet.* **4**, 2259–2264
- Stassen, M. J., Bailey, D., Nelson, S., Chinwalla, V., and Harte, P. J. (1995) *Mech. Dev.* **52**, 209–223
- Freedman, D. A., and Levine, A. J. (1998) *Mol. Cell. Biol.* **18**, 7288–7293
- Roth, J., Dobbelsstein, M., Freedman, D. A., Shenk, T., and Levine, A. J. (1998) *EMBO J.* **17**, 554–564
- Haupt, Y., Maya, R., Kazanietz, A., and Oren, M. (1997) *Nature* **387**, 296–299
- Kubbutat, M. H., Jones, S. N., and Vousden, K. H. (1997) *Nature* **387**, 299–303

<sup>2</sup> I. A. Grigorian, K. V. Gurova, M. Zeremski, S. S. Kwek, and A. V. Gudkov, manuscript in preparation.